

# Linear Time Algorithm for Optimal Feed-link Placement

Marko Savić <sup>\*</sup>      Miloš Stojaković <sup>†</sup>

## Abstract

A polygon representing transportation network is given, together with a point  $p$  in its interior. We aim to extend the network by inserting a line segment, called a feed-link, which connects  $p$  to the boundary of the polygon. Geometric dilation of some point  $q$  on the boundary is the ratio between the length of the shortest path from  $p$  to  $q$  through the extended network and their Euclidean distance. The utility of a feed-link is inversely proportional to the maximal dilation over all boundary points. We give a linear time algorithm for computing the feed-link with the minimum overall dilation, thus improving upon the previously known algorithm of complexity close to  $O(n \log n)$ .

## 1 Introduction

Given a polygon and a point  $p$  inside it, we want to connect  $p$  to the boundary of the polygon using a single line segment. Any such connection is called a *feed-link*. The optimal feed-link is the one that minimizes detour factor from point  $p$  to any other point on the boundary.

The problem of finding the optimal feed-link is solved in [1] in  $O(\lambda_7(n) \log n)$  time, where  $\lambda_7(n)$  is a slightly superlinear function. We make an improvement by presenting an  $O(n)$  time algorithm.

Although the problem statement assumes that  $p$  lies inside the polygon, all the calculation work out exactly the same for an arbitrary point  $p$  in the plane, and the same result is obtained in that more general setting.

## 2 Notation and problem statement

Polygon with boundary  $P$  is defined by vertices  $v_0, v_1, \dots, v_{n-1}$ . The point  $p$  in its interior is given.

*Feed-link* is any line segment  $pq$ , connecting  $p$  with some point  $q \in P$  on the boundary.

---

<sup>\*</sup>Department of Mathematics and Informatics, University of Novi Sad, Serbia. Email: marko.savic@dmi.uns.ac.rs. Partly supported by Ministry of Education and Science, Republic of Serbia

<sup>†</sup>Department of Mathematics and Informatics, University of Novi Sad, Serbia. Email: milos.stojakovic@dmi.uns.ac.rs. Partly supported by Ministry of Education and Science, Republic of Serbia, and Provincial Secretariat for Science, Province of Vojvodina.

Given points  $q, r \in P$  *dilation of  $r$  via  $q$*  is defined as

$$\delta_q(r) = \frac{|pq| + \text{dist}(q, r)}{|pr|},$$

where  $\text{dist}(q, r)$  is a distance between  $q$  and  $r$  by going over the polygon's boundary, and  $|ab|$  is the Euclidean distance between points  $a$  and  $b$ , see Figure 1.

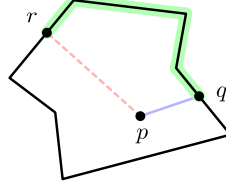


Figure 1: The concept of dilation.

Given point  $q \in P$ , *dilation via  $q$*  is defined as

$$\delta_q = \max_{r \in P} \delta_q(r).$$

The problem of finding the optimal feed-link is to find  $q$  such that  $\delta_q$  is minimized.

## 2.1 Left and right dilation

We split the concept of dilation into the *left* and *right* dilation.

Given two points  $a, b \in P$ ,  $P[a, b]$  is the portion of the  $P$  obtained by going from  $a$  to  $b$  around the polygon in the positive direction, including points  $a$  and  $b$ . Let  $\mu(a, b)$  be the length of  $P[a, b]$  and  $\mu(P)$  is the length (perimeter) of  $P$ .

For given  $a \in P$ , let  $a'$  be the point on  $P$ , different from  $a$ , for which  $\mu(a, a') = \mu(a', a) = \mu(P)/2$ , see Figure 2. By  $P^+[a]$  we denote  $P[a, a']$ , and by  $P^-[a]$  we denote  $P[a', a]$ . Obviously,  $P^+[a] \cup P^-[a] = P$  and  $P^+[a] \cap P^-[a] = \{a, a'\}$ .

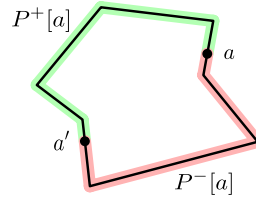


Figure 2: Left and right portion of  $P$  from the point  $a$ .

Given point  $q \in P$  and  $r \in P^+[q]$ , *left dilation of  $r$  via  $q$*  is defined as

$$\delta_q^+(r) = \frac{|pq| + \mu(q, r)}{|pr|}.$$

Given point  $q \in P$  and  $r \in P^-[q]$ , *right dilation of  $r$  via  $q$*  is defined as

$$\delta_q^-(r) = \frac{|pq| + \mu(r, q)}{|pr|}.$$

When measuring  $\text{dist}(q, r)$  we take the shortest path from  $q$  to  $r$  over  $P$ , and that whole path must lie entirely either in  $P^+[q]$  or  $P^-[q]$ . This allows us to express the dilation of  $r$  via  $q$  using left and right dilations of  $r$  via  $q$

$$\delta_q(r) = \begin{cases} \delta_q^+(r) & \text{if } r \in P^+[q] \\ \delta_q^-(r) & \text{if } r \in P^-[q] \end{cases}.$$

Given point  $q \in P$ , *left dilation via  $q$*  is defined as  $\delta_q^+ = \max_{r \in P^+[q]} \delta_q^+(r)$ , and *right dilation via  $q$*  as  $\delta_q^- = \max_{r \in P^-[q]} \delta_q^-(r)$ . Finally, the dilation via  $q$  can be expressed as

$$\delta_q = \max(\delta_q^+, \delta_q^-) = \max_{r \in P} \delta_q(r). \quad (1)$$

In the next two sections we will be concerned only about left dilations; analogous result will hold for right dilation. In Section 5 we will show how to combine our findings about left and right dilation to compute the answer to the stated question. Here we stop introducing notations with  $+$  and  $-$  in superscripts denoting left and right dilations, and consider each new notation to be related with left dilation, if applicable.

### 3 Another view of the problem

We parametrize points on  $P$  by defining  $P(t)$ ,  $t \in \mathbb{R}$ , to be the point on  $P$  for which  $\mu(v_0, P(t)) \equiv t \pmod{\mu(P)}$ , see Figure 3.

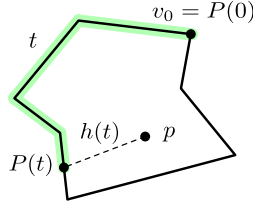


Figure 3: Parametrization of  $P$ .

The plot of the function  $h(t) = |pP(t)|$  is an infinite sequence of hyperbola segments joined at their endpoints, where  $(kn + r)$ -th hyperbola segment corresponds to the  $r$ -th side of  $P$ , for  $r \in \{0, 1, \dots, n-1\}$  and  $k \in \mathbb{Z}$ , see Figure 4. For each  $i = kn + r$ , hyperbola  $h_i$  is of the form  $h_i(t) = \sqrt{(t - m_i)^2 + d_i^2}$ , where  $m_{kn+r} = m_r + k\mu(P)$ , and  $d_{kn+r} = d_r$ ,  $d_r$  being the distance between  $p$  and the line containing  $r$ -th side of  $P$ . The left endpoint of  $i$ -th hyperbola segment is  $E_i = (e_i, h(e_i))$ , where  $e_i = k\mu(P) + \mu(v_r)$ , and the right endpoint is at  $(e_{i+1}, h(e_{i+1}))$ , so  $e_i < e_{i+1}$ . We will consider that each hyperbola segment contains its left endpoint, but not the right endpoint. Let  $H(t) = (t, h(t))$ . We denote  $i$ -th hyperbola segment with  $\mathcal{H}_i$ . The plot is, obviously, periodic, with the period of  $\mu(P)$ , that is,  $H(t) = H(t + k\mu(P))$ .

Let  $o(t) = t - h(t)$ , and  $O(t) = (o(t), 0)$ . We also define  $o_i(t) = t - h_i(t)$ , and  $O_i(t) = (o_i(t), 0)$ . For given  $t_q$  and  $t_r$ , such that  $t_q \leq t_r \leq t_q + \mu(P)/2$ , the slope of the line passing through points  $O(t_q)$  and  $H(t_r)$  is

$$s(t_q, t_r) = \frac{h(t_r)}{t_r - t_q + h(t_q)} = \frac{|pP(t_r)|}{\mu(P(t_q), P(t_r)) + |pP(t_q)|} = \frac{1}{\delta_{P(t_q)}^+(P(t_r))},$$

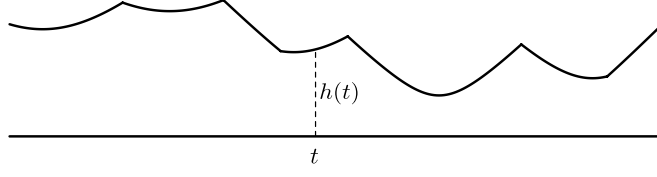


Figure 4: The plot of  $h(t)$ .

hence the slope between  $O(t_q)$  and  $H(t_r)$  equals the inverse of the left dilation of  $P(t_r)$  via  $P(t_q)$ , see Figure 5.

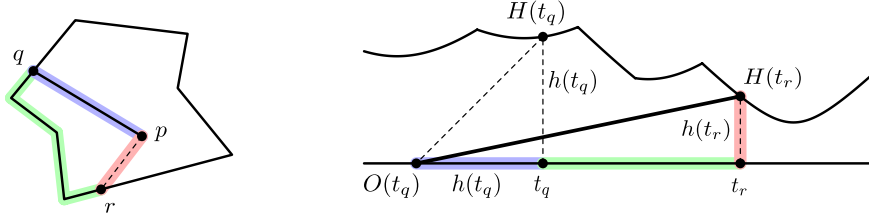


Figure 5: Dilation and slope relation.

The lowest slope  $s(t_q)$  from  $O(t_q)$  to  $H(t_r)$  among all  $t_r \in [t_q, t_q + \mu(P)/2]$  thus equals the inverse of the left dilation via  $P(t_q)$ :

$$\begin{aligned} s(t_q) &= \min_{t_r \in [t_q, t_q + \mu(P)/2]} s(t_q, t_r) = \min_{t_r \in [t_q, t_q + \mu(P)/2]} \frac{1}{\delta_{P(t_q)}^+(P(t_r))} \\ &= \frac{1}{\max_{r \in P^+[P(t_q)]} \delta_{P(t_q)}^+(r)} = \frac{1}{\delta_{P(t_q)}^+}. \end{aligned} \quad (2)$$

Obviously,  $s(t) \in (0, 1]$  because it is strictly positive and  $s(t) \leq s(t, t) = 1$ .

**Lemma 1.** *For any two distinct values of  $t_1$  and  $t_2$ ,*

$$|(h(t_2) - h(t_1))/(t_2 - t_1)| \leq 1.$$

*Proof.* Function  $h(t)$  is continuous and, as a union of countably many segments of hyperbolas with first derivatives less or equal one in absolute value, is differentiable almost everywhere having  $|h'(t)| < 1$  for each  $t \in \mathbb{R} \setminus \{e_i : i \in \mathbb{N}\}$ .  $\square$

So far,  $s(t_q)$  was defined as minimum only among slopes  $s(t_q, t_r)$  where  $t_r \in [t_q, t_q + \mu(P)/2]$ . However, this range can be extended since from Lemma 1 follows that  $s(t_q, t_r)$  cannot be less than 1 when  $t_r \in [o(t_q), t_q]$ . Hence,

$$s(t_q) = \min_{t_r \in [o(t_q), t_q + \mu(P)/2]} s(t_q, t_r).$$

## 4 Sliding lever algorithm

Consider the line segment having slope  $s(t)$ , with one endpoint at  $O(t)$  and the other at  $(t + \mu(P)/2, s(t)(\mu(P)/2 + h(t)))$ . Let us call that line segment the *lever for  $t$*  (Figure 6). Note that the lever only touches the plot, never intersecting it properly.

Let  $C(t) = H(c(t)) = (c(t), h(c(t)))$  be the leftmost point in which the lever for  $t$  touches the plot. Then,  $c(t) \in [o(t), t + \mu(P)/2]$  is the lowest value for which  $s(t) = s(t, c(t))$ . From the original perspective, this means that left dilation via  $P(t)$  reaches maximum at  $P(c(t))$ .

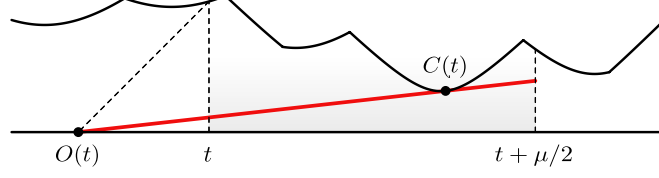


Figure 6: Lever.

We continuously decrease parameter  $t$  and observe what is happening with the lever for  $t$ . Decreasing  $t$  corresponds to “dragging” the lever strictly in the leftward direction, because when  $t$  is decreasing, the following monotonicity lemma states that  $o(t)$  and  $c(t)$  are decreasing as well.

**Lemma 2.**  $t_1 < t_2$  implies  $o(t_1) \leq o(t_2)$  and  $c(t_1) \leq c(t_2)$

*Proof.* Suppose  $t_1 < t_2$ . From Lemma 1:

$$\begin{aligned} (h(t_2) - h(t_1)) / (t_2 - t_1) &\leq 1 \\ t_1 - h(t_1) &\leq t_2 - h(t_2) \\ o(t_1) &\leq o(t_2). \end{aligned}$$

To show that  $c(t_1) \leq c(t_2)$ , assume the opposite, that  $t_1 < t_2$  and  $c(t_1) > c(t_2)$ . Then,  $t_1 < t_2 < c(t_2) < c(t_1) \leq t_1 + \mu(P)/2$ , and  $c(t_1) \in [t_2, t_2 + \mu(P)/2]$ .

If line segments  $O(t_1)C(t_1)$  and  $O(t_2)C(t_2)$  do not intersect, then  $O(t_2)C(t_2)$  lies completely under the  $O(t_1)C(t_1)$  and, since  $O(t_2)C(t_2)$  touches the plot, the plot must intersect  $O(t_1)C(t_1)$  in some point left of  $t_1$ , which is impossible since  $C(t_1)$  is the leftmost point where the lever for  $t_1$  touches the plot.

Otherwise, suppose  $O(t_1)C(t_1)$  and  $O(t_2)C(t_2)$  intersect.

$$s(t_2, c(t_1)) < s(t_2, c(t_2)) = \min_{t \in [t_2, t_2 + \mu(P)/2]} s(t_2, t) \leq s(t_2, c(t_1)),$$

which is a contradiction. □

We now introduce some functions to be used later.

Let  $C_j(o) = H(c_j(o)) = (c_j(o), h(c_j(o)))$  be the contact point of hyperbola  $h_j$  and its tangent through the point  $(o, 0)$ ,  $o < m_j$ . It can be calculated by solving the equation  $h'_j(c_j(o)) = h_j(c_j(o)) / (c_j(o) - o)$ , which results in

$$c_j(o) = \frac{d_j^2 + m_j^2 - m_j o}{m_j - o}. \quad (3)$$

For example, if we know that the lever for  $t$  is tangent to the hyperbola segment  $\mathcal{H}_j$ , we then know that it is touching it at the point  $c_j(o(t))$ .

Next, let  $s_j(o)$  be the slope of that tangent, the line through  $(o, 0)$  and  $C_j(o)$ .

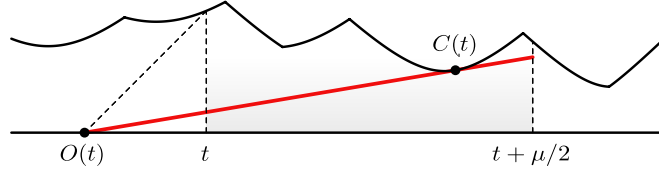
$$s_j(o) = \frac{h_j(c_j(o))}{c_j(o) - o} = 1/\sqrt{\left(\frac{m_j - o}{d_j}\right)^2 + 1}. \quad (4)$$

#### 4.1 States

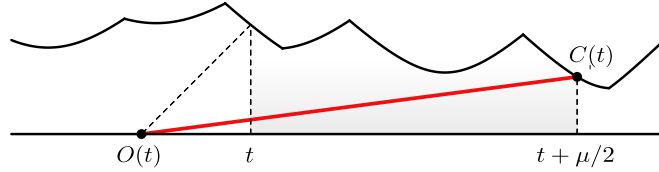
We define different lever states depending on where  $t$  and  $c(t)$  are.

When  $t \in [e_i, e_{i+1})$  and  $c(t) \in [e_j, e_{j+1})$ , we say that the lever for  $t$  is in state  $\langle i, j \rangle$ . There are three possible types of positions defined by the manner in which the lever for  $t$  touches the plot:

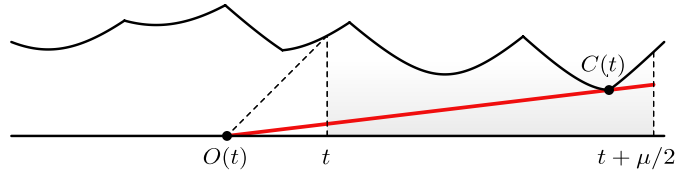
- $\langle i, j \rangle^{\mathcal{K}}$  :  $c(t) < t + \mu(P)/2$  and the lever is the tangent to  $\mathcal{H}_j$ . The lever is sliding to the left along the  $h_j$  maintaining the tangency, thus continuously decreasing the slope.



- $\langle i, j \rangle^{\mathcal{Y}}$  :  $c(t) = t + \mu(P)/2$ . Point  $C(t)$  is the right endpoint of the lever. It is the only point where lever touches the plot. The lever is moving to the left while keeping its right endpoint on  $h_j$ .



- $\langle i, j \rangle^{\mathcal{V}}$  :  $c(t) < t + \mu(P)/2$  and the lever is passing through the point  $H(e_j)$ , the endpoint between hyperbola segments  $\mathcal{H}_{j-1}$  and  $\mathcal{H}_j$ . This situation occurs only if  $m_{j-1} > m_j$ . The two neighbouring hyperbola segments then form a “wedge” pointing downwards, and the lever is sliding to the left while maintaining the contact with the tip of that wedge, thus continuously decreasing the slope.



## 4.2 Events

During the process of dragging the lever towards the left, the lever state changes at certain moments. We call such events *state transition events*. If the current value of  $t$ , denoted by  $t_c$ , and the current state are known, the following event can be determined by maintaining the set of conceivable future events of which at least one must be realized, and proceeding to the one that is first to happen, i.e. the one with the largest  $t$  among them. That is why we must be able to calculate the value of  $t$  for each of those events.

For all events we will give a polynomial equation describing it. The equations will be solved either for  $t$  or  $o_i(t)$ . Once we have  $o_i(t)$  we can easily obtain  $t$  from the following inverse function of  $o_i(t)$ .

$$t = \frac{o_i(t)^2 - d_i^2 - m_i^2}{2(o_i(t) - m_i)}. \quad (5)$$

In the process of determining  $t$ , we will repeatedly encounter fixed degree polynomial equations. Solving them can be assumed to be a constant time operation, see [2].

### 4.2.1 Jumping and retargeting

*Jumping* is any event in which  $C(t)$  abruptly changes its position by switching to some other hyperbola segment. In order to efficiently find state transition events that include jumps, we always need to know which hyperbola segment can be jumped to from the current position. There is always at most one such target hyperbola segment, and we will see how to keep track of it.

Consider some point  $H(x)$  on the plot. Let  $\text{jump}(x)$ , the *jump destination* for  $x$ , be the index of a hyperbola segment which contains the rightmost point  $H(w)$  on the plot such that  $w < x$  and the ray from  $H(x)$  through  $H(w)$  only touches the plot, but does not intersect it properly. That is,  $\text{jump}(x)$  is the index of the lowest visible hyperbola segment when looking from the point  $H(x)$  to the left. If there is no such  $w$ , because hyperbola segments on the left are obscured by the one containing  $H(x)$ , then we set  $\text{jump}(x)$  to be the index of the hyperbola segment containing  $H(x)$ .

Consider only the values of  $x$  where  $\text{jump}(x)$  changes. We call such values *retargeting positions*, and points  $H(x)$  *retargeting points* (Figure 7).

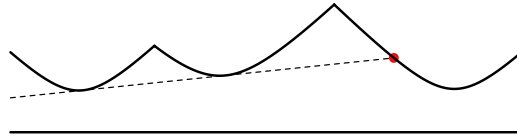


Figure 7: Retargeting point.

**Theorem 1.** *It is possible to find all retargeting positions, ordered from left to right, together with jump destinations of those positions in  $O(n)$  time.*

*Proof.* The algorithm for finding all retargeting points is similar to finding the lower convex chain in Andrew's monotone chain convex hull algorithm [3], except the sorting part is not needed.

Let  $j_0$  be the index of the hyperbola segment which contains any of the lowest points of the plot of  $h$ . We scan segments from left to right and maintain the longest chain of those hyperbola segments that are in a convex position, starting with the  $\mathcal{H}_{j_0}$  and ending with the segment currently being processed. Intuitively, convex position means that for each three hyperbola segments appearing consequently in the convex chain, left and right of those segments cannot be connected with a line segment passing completely under the middle one, see Figure 8.

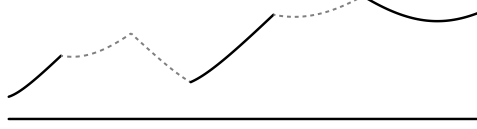


Figure 8: Hyperbola segments drawn with a solid line are in convex position. Remember that each hyperbola segment contains its left endpoint, but not the right one.

The stack is used for keeping the track of segments in the chain (we store their indices only). At each of the following segments encountered, we pop segments from the stack until the current segment can be added to the chain so that the chain is again in convex position, after which we push the current segment onto stack.

---

**Algorithm 1** Retargeting Points

---

```

RetargetingPoints  $\leftarrow$  [ ]
PUSH( $j_0 - 1$ )
PUSH( $j_0$ )
for  $k \leftarrow j_0 + 1$  to  $j_0 + n$  do
  loop
     $i \leftarrow$  Stack second
     $j \leftarrow$  Stack top
     $l_1 \leftarrow$  TANGENT( $i, j$ )
     $l_2 \leftarrow$  TANGENT( $j, k$ )
    if  $l_1 \cap \mathcal{H}_k \neq \emptyset$  then
      Let  $g$  be the leftmost point from  $l_1 \cap \mathcal{H}_k$ .
      Append  $g$  to RetargetingPoints, and set jump destination of  $g$  to  $j$ .
      POP()
    else
      if  $l_2 \cap \mathcal{H}_k \neq \emptyset$  then
        Let  $g$  be the leftmost point from  $l_2 \cap \mathcal{H}_k$ .
        Append  $g$  to RetargetingPoints, and set jump destination of  $g$  to  $j$ .
        PUSH( $k$ )
      end if
      break loop
    end if
  end loop
end for

```

---

TANGENT( $i, j$ ) returns the line which acts as a boundary between the region of the plane from which the  $\mathcal{H}_i$  is visible from below of  $\mathcal{H}_j$  and the region from which  $\mathcal{H}_i$  is obscured by the  $\mathcal{H}_j$ . More precisely, TANGENT( $i, j$ ) is the line touching both  $\text{cl}(\mathcal{H}_i)$  and  $\text{cl}(\mathcal{H}_j)$  from below, where  $\text{cl}(X)$  denotes the closure of a point set  $X$ , which is, in our case, essentially an addition of the right endpoint to the hyperbola segment.



Each point reported is indeed a retargeting point since the jump destination changes at it. Points reported in the “if” branch have the property that points on  $\mathcal{H}_k$  immediately left and right from  $g$  have  $\mathcal{H}_j$  and  $\mathcal{H}_i$  as their jump destinations, respectively. Points reported in the “else” branch have the property that points on  $\mathcal{H}_k$  immediately left and right from  $g$  have  $\mathcal{H}_j$  and  $\mathcal{H}_k$  as their jump destinations, respectively.

Let us make sure that no retargeting points were omitted. Each retargeting point  $g$  on  $\mathcal{H}_k$  must lie either on  $\text{TANGENT}(i, j)$  or  $\text{TANGENT}(j, k)$ , for some  $i < j < k$ . In the first case,  $\mathcal{H}_i$  and  $\mathcal{H}_j$  must be two consecutive elements of the convex chain ending with  $\mathcal{H}_{k-1}$ , otherwise  $\mathcal{H}_i$  or  $\mathcal{H}_j$  would not be visible from the neighbourhood of  $g$ . Since  $\text{TANGENT}(i, j)$  intersects  $\mathcal{H}_k$ , the same must be true for each subsequent pair from the convex chain on the right of  $i$  and  $j$ . Thus, while the algorithm pops elements from the stack, in one moment the last two elements on the stack will happen to be exactly  $i$  and  $j$ . At that iteration of the inner loop  $g$  will be reported in the “if” branch as the retargeting point. The other case, when  $g$  is on the  $\text{TANGENT}(j, k)$  is similar:  $\mathcal{H}_j$  must be the element of the convex chain ending with  $\mathcal{H}_{k-1}$ , otherwise it would not be visible from the neighbourhood of  $g$ . The tangent of each subsequent pair from the convex chain with the right element on the right of  $j$  intersects  $\mathcal{H}_k$ . Thus, the algorithm will be popping elements from the stack until the top element of the stack is  $j$ , when  $g$  will be reported in the “else” branch as the retargeting point.

The running time of algorithm is  $O(n)$ , since  $\text{TANGENT}()$  can be computed in constant time, and each hyperbola segment index  $k \in \{j_0 + 1, \dots, j_0 + n\}$  is pushed on stack and popped from stack at most once. Hence, the number of retargeting points reported is also  $O(n)$ .

Retargeting points reported by the algorithm come in order sorted from left to right. This is because the retargeting points reported in a single iteration of the outer for-loop belong to the same hyperbola segment, and segments come in left-to-right order. Retargeting points reported on the same hyperbola segment are also in the left-to-right order: inside the inner loop,  $\mathcal{H}_k$  is consecutively intersected with lines such that each line is of the lower slope than previous and lies beneath it under  $\mathcal{H}_k$ . Hence, each subsequent intersection point lies to the right of the previous one.

Here we reported only retargeting points from a single period of the plotted function, but all other can be found by simply translating these horizontally by the integral number of periods  $\mu(P)$ .  $\square$

Next, we list all types of events that can happen while moving the lever leftwards and show how to calculate the value of  $t$  for each. When talking about events that involve jumps, the current jump destination,  $\text{jump}(c(t))$ , will be denoted by  $j_m$ .

#### 4.2.2 Jump destination change event

Jump destination  $j_m$  changes whenever  $C(t)$  passes over some retargeting point. At that moment it is necessary to recalculate all future events which involve jumps. Let  $z$  be the next retargeting position, i.e. the rightmost one that lie to the left of  $t_c$ . Depending on the lever state, we calculate the event position in one of the following ways.

- If the current state is  $\langle i, j \rangle^{\mathcal{V}}$  – The jump destination change event cannot occur before leaving this state since  $C(t)$  stands still at the “wedge tip”, so it cannot pass any other point.
- If the current state is  $\langle i, j \rangle^{\mathcal{K}}$  – The lever is tangent to the  $\mathcal{H}_j$ , so this event can only happen if  $z > m_j$ . Otherwise, the lever would have nonpositive slope when touching the plot at  $H(z)$ . The equation describing this event is

$$c_j(o_i(t)) = z,$$

and it solves to

$$o_i(t) = \frac{d_j^2 - zm_j + m_j^2}{m_j - z}.$$

- If the current state is  $\langle i, j \rangle^{\mathcal{Y}}$  – Right endpoint of the lever slides over  $\mathcal{H}_j$  and will coincide with  $H(z)$  when

$$t = z - \mu(P)/2.$$

Note that  $j_m$  is not used to describe lever state, so jump destination change event does not change the current state. All other considered events, however, are state transition events.

#### 4.2.3 $\langle i, j \rangle \rightarrow \langle i - 1, j \rangle$

The interval to which  $t$  belongs changes from  $[e_i, e_{i+1})$  to  $[e_{i-1}, e_i)$ . This event happens at  $e_i$ .

$$t = e_i.$$

#### 4.2.4 $\langle i, j \rangle^{\mathcal{Y}} \rightarrow \langle i, j - 1 \rangle^{\mathcal{Y}}$

This is the event when the right endpoint of the lever simply slides from one hyperbola segment to another.

$$t = e_j - \mu(P)/2.$$

#### 4.2.5 $\langle i, j \rangle^{\mathcal{Y}} \rightarrow \langle i, j \rangle^{\mathcal{K}}$ and $\langle i, j \rangle^{\mathcal{K}} \rightarrow \langle i, j \rangle^{\mathcal{Y}}$

In this event the lever changes from being a tangent to  $\mathcal{H}_j$  to touching  $\mathcal{H}_j$  with its right endpoint, or the other way round. The corresponding equation for this event is

$$c_j(o_i(t)) = t + \mu(P)/2,$$

which can be transformed to a cubic equation in  $t$ . Since there can be at most three real solutions to that equation, it is possible that this event takes place at most three times with the same  $i$  and  $j$ . On each occurrence of the event the lever switches between being a tangent and touching the plot with its right endpoint.

#### 4.2.6 $\langle i, j \rangle^{\mathcal{V}} \rightarrow \langle i, j_m \rangle^{\mathcal{K}}$

This event happens when the lever state changes from having an endpoint on  $\mathcal{H}_j$  to being a tangent to  $\mathcal{H}_{j_m}$ . The equation is

$$s_{j_m}(o_i(t)) = \frac{h_j(t + \mu(P)/2)}{h_i(t) + \mu(P)/2},$$

which further transforms into a 5th order polynomial equation in  $t$ .

The line through  $o_i(t)$  with slope  $s_{j_m}(o_i(t))$  touches the hyperbola  $h_{j_m}$ , but we need to be sure that it actually touches the segment  $\mathcal{H}_{j_m}$  of that hyperbola. It may as well be the case that  $\mathcal{H}_{j_m}$  is not wide enough to have a common point with the line. More precisely, the first coordinate,  $u$ , of the touching point between the line and  $h_{j_m}$  must belong to interval  $[e_{j_m}, e_{j_m+1})$ . To get that coordinate, we solve the equation

$$h'_{j_m}(u) = s_{j_m}(o_i(t)).$$

Considering that  $o_i(t) < m_{j_m} < u$  must hold, we get a single solution

$$u = m_{j_m} + \frac{d_{j_m}^2}{m_{j_m} - o_i(t)}.$$

If  $u \notin [e_{j_m}, e_{j_m+1})$ , we do not consider this event.

This check, to see if the line through  $o_i(t)$  with slope  $s_{j_m}(o_i(t))$  actually touches the hyperbola segment  $\mathcal{H}_{j_m}$ , we call *collision check*, and we will also use it in some other events.

#### 4.2.7 $\langle i, j \rangle^{\mathcal{V}} \rightarrow \langle i, j_m \rangle^{\mathcal{V}}$

The event when the lever state changes from having an endpoint on  $\mathcal{H}_j$  to touching the wedge between  $\mathcal{H}_{j_{m-1}}$  and  $\mathcal{H}_{j_m}$  is described by

$$\frac{h_{j_m}(e_{j_m})}{e_{j_m} - o_i(t)} = \frac{h_j(t + \mu(P)/2)}{h_i(t) + \mu(P)/2},$$

which again transforms into a 5th order polynomial equation in  $t$ .

#### 4.2.8 $\langle i, j \rangle^{\mathcal{K}} \rightarrow \langle i, j \rangle^{\mathcal{V}}$

This event happens when the point in which the lever is touching  $\mathcal{H}_j$  reaches  $e_j$ . Here, the lever is tangent to  $\mathcal{H}_j$ , and since it must have a positive slope, this will only happen if  $e_j > m_j$ . The event equation is

$$c_j(o_i(t)) = e_j,$$

which solves to

$$o_i(t) = \frac{d_j^2 - e_j m_j + m_j^2}{m_j - e_j}.$$

#### 4.2.9 $\langle i, j \rangle^K \rightarrow \langle i, j_m \rangle^K$

This event happens when the lever becomes a tangent to two hyperbola segments,  $\mathcal{H}_j$  and  $\mathcal{H}_{j_m}$  simultaneously. It can only happen if  $\mathcal{H}_{j_m}$  is lower than  $\mathcal{H}_j$ , i.e.  $d_{j_m} < d_j$ .

$$s_{j_m}(o_i(t)) = s_j(o_i(t)).$$

Since  $o_i(t) < m_{j_m}$  and  $o_i(t) < m_j$ , the only solution is

$$o_i(t) = \frac{d_j m_{j_m} - d_{j_m} m_j}{d_j - d_{j_m}}.$$

Here we need to apply the collision check explained in Section 4.2.6 to see if the common tangent actually touches  $\mathcal{H}_{j_m}$ . If the test fails, we do not consider this event.

#### 4.2.10 $\langle i, j \rangle^K \rightarrow \langle i, j_m \rangle^V$

The event in which the lever touches the wedge tip at  $e_{j_m}$  while being a tangent to  $h_j$  is represented by the following equation. This can only happen if  $h_{j_m}(e_{j_m}) < d_j$ .

$$\frac{h_{j_m}(e_{j_m})}{e_{j_m} - o_i(t)} = s_j(o_i(t))$$

This can be transformed to a quadratic equation in  $o_i(t)$ . The two solutions correspond to two tangents to  $h_j$  from the point  $(e_{j_m}, h(e_{j_m}))$ . The smaller of the two solutions is where this event happens.

#### 4.2.11 $\langle i, j \rangle^V \rightarrow \langle i, j - 1 \rangle^V$

This event happens when the lever stops touching the tip of the wedge and starts to slide its right endpoint over the hyperbola segment on the left of the wedge.

$$t = e_j - \mu(P)/2.$$

#### 4.2.12 $\langle i, j \rangle^V \rightarrow \langle i, j - 1 \rangle^K$

This event happens when the lever stops touching the tip of the wedge and becomes a tangent of the hyperbola segment on the left of the wedge. This can only happen if  $e_j > m_{j-1}$ .

$$c_{j-1}(o_i(t)) = e_j,$$

which solves to

$$o_i(t) = \frac{d_{j-1}^2 - e_j m_{j-1} + m_{j-1}^2}{m_{j-1} - e_j}.$$

#### 4.2.13 $\langle i, j \rangle^{\mathcal{V}} \rightarrow \langle i, j_m \rangle^{\mathcal{K}}$

This event happens when the lever stops touching the tip of the wedge and becomes a tangent of the hyperbola segment  $\mathcal{H}_{j_m}$ . This can only happen if  $h_j(e_j) > d_{j_m}$ .

$$s_{j_m}(o_i(t)) = \frac{h_j(e_j)}{e_j - o_i(t)}$$

From that we get a quadratic equation in  $o_i(t)$ . The two solutions correspond to two tangents to  $h_{j_m}$  from the point  $(e_j, h(e_j))$ . The smaller of the two solutions is where this event happens.

#### 4.2.14 $\langle i, j \rangle^{\mathcal{V}} \rightarrow \langle i, j_m \rangle^{\mathcal{V}}$

This event happens when the lever touches two wedges, at points  $E_{j_m}$  and  $E_j$  simultaneously. The condition for that is

$$\frac{h_j(e_j)}{e_j - o_i(t)} = \frac{h_{j_m}(e_{j_m})}{e_{j_m} - o_i(t)},$$

which is a simple linear equation in  $o_i(t)$ .

#### 4.2.15 Sequence of states

We want to find a sequence of states through which the lever will pass on its leftward journey, together with the positions where state changes happen. Let the obtained sequence be  $p_1, \mathcal{S}_1, p_2, \mathcal{S}_2, p_3, \dots, p_r, \mathcal{S}_r$ , where  $p_1 \leq p_2 \leq \dots \leq p_r$ . Each state  $\mathcal{S}_k$  takes part when the lever position is exactly between  $p_k$  and  $p_{k+1}$ , where  $p_{r+1} = p_1 + \mu(P)$ . We call this *the sequence of realized states*.

To calculate the sequence of realized states, we start from a specific lever position that has a known state. Let  $p_{low}$  be any of the values for which  $h(p_{low})$  is minimal, and  $\mathcal{H}_{j_0}$  be the hyperbola segment containing it. The algorithm starts with a lever in position  $t_c = t_0 = p_{low} - \mu(P)/2$ . This lever has its right endpoint on the plot at the point  $H(p_{low})$ , which means that its state is  $\langle i_0, j_0 \rangle^{\mathcal{V}}$ , where  $i_0$  is the index of hyperbola segment over  $t_0$ .  $p_{low}$  is also a retargeting position, so we also know initial jump destination.

The algorithm then iterates with the following operations in its main loop. It first calculates all possible events that could happen while in the current state. Among those events let  $E$  be the one with the largest  $t$  that is not larger than current  $t_c$ . It is the event that must occur next. The algorithm sets  $t_c$  to equal  $t$ , and updates jump destination if  $E$  is jump destination change event, or switches to the new state if the event is state transition event. In the latter case, position  $t$  and the new state are added to the sequence of realized states. These operations are iterated until one whole period of the plot is swept, ending with  $t_c = t_0 - \mu(P)/2$  in state  $\langle i_0 - n, j_0 - n \rangle^{\mathcal{V}}$ .

**Theorem 2.** SLIDING LEVER ALGORITHM runs in  $O(n)$  time, and the length of the produced sequence of realized states is  $O(n)$ .

---

**Algorithm 2** Sliding Lever Algorithm

---

```
Find  $p_{low}$  and  $j_0$ .  
 $t_0 \leftarrow p_{low} - \mu(P)/2$   
Find  $i_0$ .  
Run RETARGETING POINTS to find retargeting points and their jump destinations.  
 $t_c \leftarrow t_0$   
 $i \leftarrow i_0$   
 $j \leftarrow j_0$   
 $j_m \leftarrow$  jump destination of  $p_{low}$   
Set the current state to  $\langle i_0, j_0 \rangle^{\mathcal{Y}}$ .  
Add  $t_0$  and  $\langle i_0, j_0 \rangle^{\mathcal{Y}}$  to the sequence.  
while  $t_c > t_0 - \mu(P)$  do  
    Calculate all the events for the current state. Ignore jumping events if  $j = j_m$ .  
    Let  $E$  be the first event to happen (the one with the largest  $t$  not larger than  $t_c$ ).  
     $t_c \leftarrow t$  of the event  $E$ .  
    if  $E$  is jump destination change event then  
        Update  $j_m$ .  
    else  
        Set the current state to the destination state of  $E$ .  
        Add  $t$  and the current state to the sequence.  
        if  $E$  is a jumping event then  
             $j_m \leftarrow j$   
        end if  
    end if  
end while
```

---

*Proof.* Each event is either jump destination change event or state transition event. From Theorem 1 we have that there are  $O(n)$  jump destination change events, and now we will show that there are  $O(n)$  state transition events.

Each state transition event transitioning from some  $\langle i, j \rangle$  state decreases either  $i$ , or  $j$  or both. The only exception are the events  $\langle i, j \rangle^{\mathcal{K}} \rightarrow \langle i, j \rangle^{\mathcal{Y}}$  and  $\langle i, j \rangle^{\mathcal{Y}} \rightarrow \langle i, j \rangle^{\mathcal{K}}$ , however those events can happen at most three times in total for the same  $i$  and  $j$ . Note that  $j_m \leq j$ , but when  $j_m = j$ , we do not consider events involving  $j_m$ . Variables  $i$  and  $j$  start with values  $i_0$  and  $j_0$ , and, after the loop finishes, they are decreased to  $i_0 - n$  and  $j_0 - n$ . Hence, no more than  $O(n)$  state transition events occurred, implying the linear length of the sequence of realized states.

Calculation of each state transition event takes a constant time, there is a constant number of events considered, and loop is iterated  $O(n)$  times. To find next jump destination change event, we move through the sorted list of retargeting positions until we find the first retargeting position not greater than  $t_c$ . The total time for calculating jump destination change events, over all iterations, is linear. Finally, the running time of the whole algorithm is also linear.  $\square$

While choosing the next event, we did not consider the possibility that there are several events with the same largest  $t$  not larger than  $t_c$ . But, if that happens, we choose an arbitrary one to be the next event. All of them will be realized in some form while lever position equals  $t$ , and no matter the order we processed them, the lever will end up in the

same state once  $t_c$  passes  $t$ .

## 5 Merging the two dilations

Knowing the sequence of realized states is sufficient to determine the exact lever slope at any position. Remember, the lever slope at position  $t$  is the inverse of the left dilation via  $P(t)$  (2). But, to know the dilation via some point we need both left and right dilations via that point (1).

Sliding lever algorithm was designed only for left dilation, but an analogous algorithm can be designed in the same manner for right dilation (or by performing the algorithm for left dilation on the mirror image of the polygon  $P$ , and then transforming obtained results appropriately). This implies the concept of the right dilation lever for  $t$  (as opposed to the left dilation lever, or just lever, as we have been calling it until now), which has negative slope and touches the plot on the left side of  $t$ . This is where we return to the original notation of  $+$  and  $-$  in superscript denoting relation with left and right dilation, respectively.

Let  $p_1^+, \mathcal{S}_1^+, p_2^+, \mathcal{S}_2^+, p_3^+, \dots, p_{r^+}^+, \mathcal{S}_{r^+}^+$  and  $p_1^-, \mathcal{S}_1^-, p_2^-, \mathcal{S}_2^-, p_3^-, \dots, p_{r^-}^-, \mathcal{S}_{r^-}^-$  be the sequences of realized states for left and right dilation, respectively, where both sequences  $p^+$  and  $p^-$  are in nondecreasing order. States for right dilation are described by  $\langle i, j \rangle$  notation as well, with the meaning analogous to the meaning of the notation for left dilation states. We say that the (right dilation) lever for  $t$  is in the state  $\langle i, j \rangle$ , when  $\mathcal{H}_i$  is the hyperbola segment above  $t$ , and the (right dilation) lever touches the hyperbola segment  $\mathcal{H}_j$ .

Having obtained both sequences of realized states, we can now merge them into a single sequence  $p_1, \mathcal{S}_1, p_2, \mathcal{S}_2, p_3, \dots, p_r, \mathcal{S}_r$ , where  $p_1 \leq p_2 \leq \dots \leq p_r$  is sorted union of  $p^+$  and  $p^-$  sequences. States of the merged sequence are pairs consisting of  $\mathcal{S}^+$  and  $\mathcal{S}^-$  state. Let  $p_{k^+}^+$  be the last position from the left dilation sequence appearing before state  $\mathcal{S}$  in the merged sequence, and correspondingly, let  $p_{k^-}^-$  be the last position from the right dilation sequence appearing before state  $\mathcal{S}$  in the merged sequence. State  $\mathcal{S}$  is then ordered pair  $(\mathcal{S}_{k^+}^+, \mathcal{S}_{k^-}^-)$ .

By the Theorem 2, both  $r^+$  and  $r^-$  are  $O(n)$ , so the length of the merged sequence is also linear in  $n$ . Hence, the merged sequence can be computed by a simple  $O(n)$  time algorithm.

For each state  $\mathcal{S}_k = (\mathcal{S}_{k^+}^+, \mathcal{S}_{k^-}^-)$  we have a single expression for computing the lever slope when  $p_k \leq t \leq p_{k+1}$ , both for the left dilation and for the right dilation. To find minimal dilation while in that state, we want to find  $t$  such that the minimum of the two slopes for the two levers for  $t$  (left and right) is maximal (1, 2)

$$\min_{p_k \leq t \leq p_{k+1}} \delta_{P(t)} = \frac{1}{\max_{p_k \leq t \leq p_{k+1}} \min\{s^+(t), s^-(t)\}}, \quad (6)$$

where  $s^+(t)$  is the same as  $s(t)$ , the slope for the left dilation lever for  $t$ , and  $s^-(t)$  is the slope for the right dilation lever for  $t$ .

Let us analyse the shape of the functions  $s^+(t)$  and  $s^-(t)$ . Assume that the corresponding state where  $t$  belongs is  $\mathcal{S} = (\mathcal{S}^+, \mathcal{S}^-)$ .

If  $\mathcal{S}^+$  is a  $\langle i, j \rangle^{\mathcal{K}}$  state, then, from equation (4), we have

$$s^+(t) = s_{\langle i, j \rangle^{\mathcal{K}}}^+(t) = s_j(o_i(t)) = \frac{h_j(c_j(o_i(t)))}{c_j(o_i(t)) - o_i(t)} = 1/\sqrt{\left(\frac{m_j - o_i(t)}{d_j}\right)^2 + 1}. \quad (7)$$

**Lemma 3.** *If  $\mathcal{S}^+$  is a  $\langle i, j \rangle^{\mathcal{K}}$  state, then  $s^+(t)$  is monotonously increasing function.*

*Proof.* From (7) we see that  $s_j(o_i(t))$  is monotonously increasing when  $o_i(t) < m_j$  which holds in the given state. Since  $o(t)$  is monotonously nondecreasing (Lemma 2), it means that their combination,  $s^+(t)$  is monotonously increasing as well.  $\square$

If  $\mathcal{S}^+$  is a  $\langle i, j \rangle^{\mathcal{V}}$  state, then we have

$$s^+(t) = s_{\langle i, j \rangle^{\mathcal{V}}}^+(t) = \frac{h_j(e_j)}{e_j - o_i(t)}. \quad (8)$$

**Lemma 4.** *If  $\mathcal{S}^+$  is a  $\langle i, j \rangle^{\mathcal{V}}$  state, then  $s^+(t)$  is monotonously increasing function.*

*Proof.* From (8) we see that  $s^+(t)$  is monotonously increasing in terms of  $o_i(t)$  when  $o_i(t) < e_j$  which holds in the given state. Since  $o(t)$  is monotonously nondecreasing (Lemma 2), it means that  $s^+(t)$  is monotonously increasing in terms of  $t$  as well.  $\square$

If  $\mathcal{S}^+$  is a  $\langle i, j \rangle^{\mathcal{Y}}$  state, then we have

$$s^+(t) = s_{\langle i, j \rangle^{\mathcal{Y}}}^+(t) = \frac{h_j(t + \mu(P)/2)}{h_i(t) + \mu(P)/2}, \quad (9)$$

which is not necessarily monotonous.

Similar observations hold for right dilation analogues:  $s^-(t)$  is monotonously decreasing if  $\mathcal{S}^-(t)$  is  $\langle i, j \rangle^{\mathcal{K}}$  or  $\langle i, j \rangle^{\mathcal{V}}$  state.

**Lemma 5.** *If  $\mathcal{S}^-$  is a  $\langle i, j \rangle^{\mathcal{Y}}$  state then  $s^+(t) \leq s^-(t)$ .*

*Proof.* From equation (2), using the fact that  $h_j(t) = h_{j+n}(t + \mu(P))$  holds because of the periodicity of the plot, we have

$$\begin{aligned} s^-(t) &= \frac{h_j(t - \mu(P)/2)}{h_i(t) + \mu(P)/2} = \frac{h_{j+n}(t + \mu(P)/2)}{h_i(t) + \mu(P)/2} = s(t, t + \mu(P)/2) \\ &\geq \min_{t_r \in [t, t + \mu(P)/2]} s(t, t_r) = s^+(t). \end{aligned}$$

$\square$

Analogously, if  $\mathcal{S}^+$  is a  $\langle i, j \rangle^{\mathcal{Y}}$  state then  $s^-(t) \leq s^+(t)$ .

We need to calculate  $\max_{p_k \leq t \leq p_{k+1}} \min\{s^+(t), s^-(t)\}$  needed in equation (6), which is equivalent of finding the highest point of the lower envelope of the functions  $s^+(t)$  and  $s^-(t)$  (Figure 9). This calculation depends on what the types of states  $\mathcal{S}^+$  and  $\mathcal{S}^-$  are.



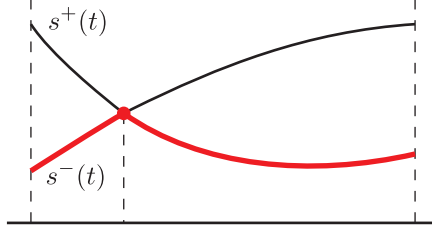


Figure 9:  $\max_{p_k \leq t \leq p_{k+1}} \min\{s^+(t), s^-(t)\}$

- $\mathcal{S}^+$  is  $\langle \cdot, \cdot \rangle^{\mathcal{Y}}$  state,  $\mathcal{S}^-$  is  $\langle \cdot, \cdot \rangle^{\mathcal{Y}}$  state : Using Lemma 5 we get  $s^+(t) = s^-(t)$ , so

$$\max_{p_k \leq t \leq p_{k+1}} \min\{s^+(t), s^-(t)\} = \max_{p_k \leq t \leq p_{k+1}} s_{\langle i, j \rangle^{\mathcal{Y}}}^+(t).$$

The maximum is achieved either at interval ends or at local maxima, which are obtained by solving a polynomial equation.

- $\mathcal{S}^+$  is  $\langle \cdot, \cdot \rangle^{\mathcal{K}}$  state,  $\mathcal{S}^-$  is  $\langle \cdot, \cdot \rangle^{\mathcal{Y}}$  state : Using Lemma 5 and Lemma 3 we get

$$\max_{p_k \leq t \leq p_{k+1}} \min\{s^+(t), s^-(t)\} = \max_{p_k \leq t \leq p_{k+1}} s_{\langle i, j \rangle^{\mathcal{K}}}^+(t) = s_{\langle i, j \rangle^{\mathcal{K}}}^+(p_{k+1}).$$

- $\mathcal{S}^+$  is  $\langle \cdot, \cdot \rangle^{\mathcal{K}}$  state,  $\mathcal{S}^-$  is  $\langle \cdot, \cdot \rangle^{\mathcal{K}}$  state : From Lemma 3 we know that  $s_{\langle i, j^+ \rangle^{\mathcal{K}}}^+(t)$  is monotonously increasing, and  $s_{\langle i, j^- \rangle^{\mathcal{K}}}^-(t)$  is monotonously decreasing. The highest point of the lower envelope of their plot on  $[p_k, p_{k+1}]$  is thus located either at one of the interval endpoints, or at the point of the intersection of two plots, which can be found by solving a polynomial equation.

Other six combinations of state types, which are not listed, are resolved in the same manner as the three mentioned above.

Finally, by taking the minimal of all  $r$  dilation minimums from  $[p_k, p_{k+1}]$  intervals we obtain the overall minimum dilation.

$$\delta = \min_{k \in \{1, 2, \dots, r\}} \min_{p_k \leq t \leq p_{k+1}} \delta_{P(t)}.$$

While calculating the minimums we maintain the value of  $t$  for which the minimum is achieved, so we also get the point on  $P$  which is the endpoint of the optimal feed-link.

## References

- [1] B. Aronov, K. Buchin, M. Buchin, B. Jansen, T. de Jong, M. van Kreveld, M. Löffler, J. Luo, R.I. Silveira, and B. Speckmann. Connect the dot: Computing feed-links for network extension. *Journal of Spatial Information Science*, (3):3–31, 2012.
- [2] A. Dickenstein and I.Z. Emiris. *Solving polynomial equations: Foundations, algorithms, and applications*, volume 14. Springer Verlag, 2005.
- [3] AM Andrew. Another efficient algorithm for convex hulls in two dimensions. *Information Processing Letters*, 9(5):216–219, 1979.